

# Relatório de Análise de Dados

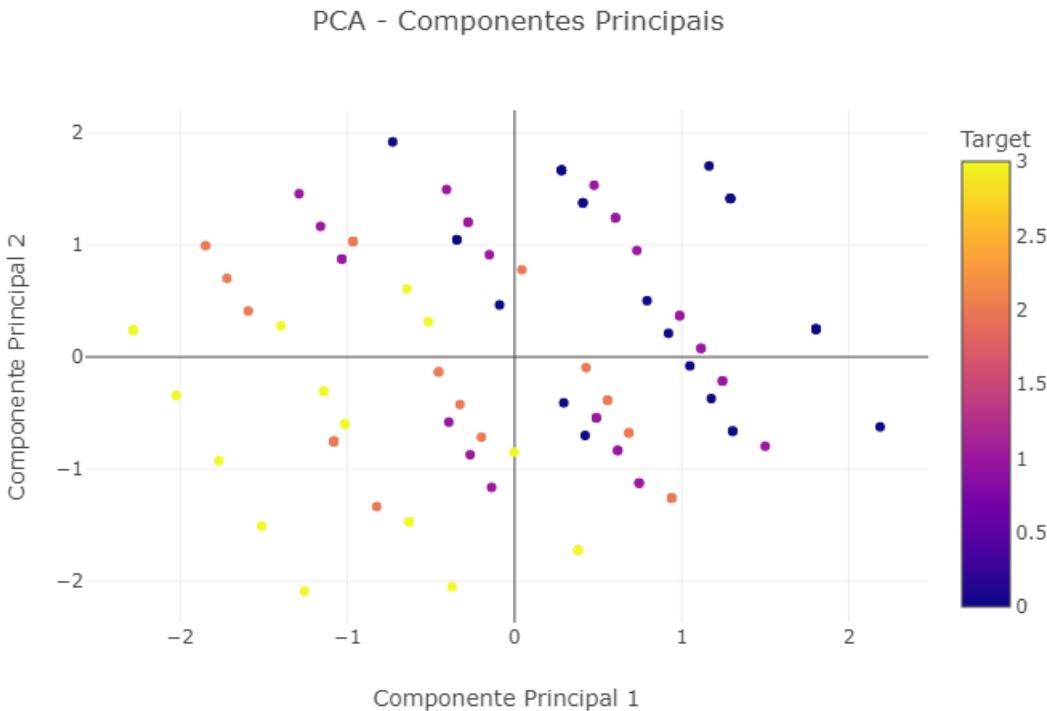
Projeto Integrador IV - UNIVESP 2024

## Objetivo

Este relatório apresenta a análise de dados e técnicas de machine learning aplicadas aos dados fornecidos. O objetivo é explorar as principais características dos dados e construir modelos preditivos usando as técnicas discutidas ao longo do curso.

## Análise de Componentes Principais (PCA)

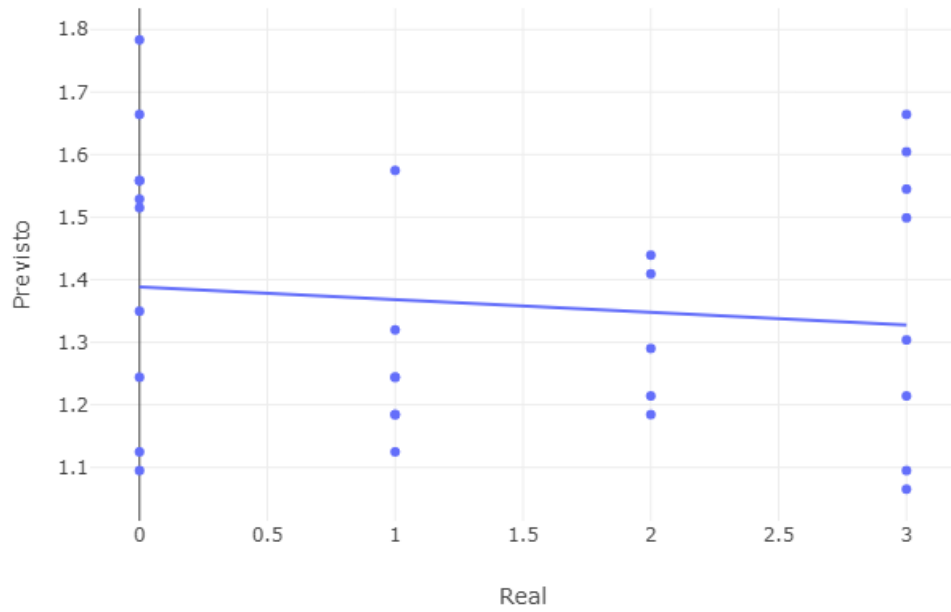
A Análise de Componentes Principais (PCA) foi realizada para reduzir a dimensionalidade dos dados e explorar as relações entre as variáveis. Abaixo está o gráfico dos componentes principais.



## Resultados da Regressão Linear

Os resultados da regressão linear indicam a qualidade do ajuste do modelo preditivo. As métricas como MSE e  $R^2$  estão incluídas abaixo.

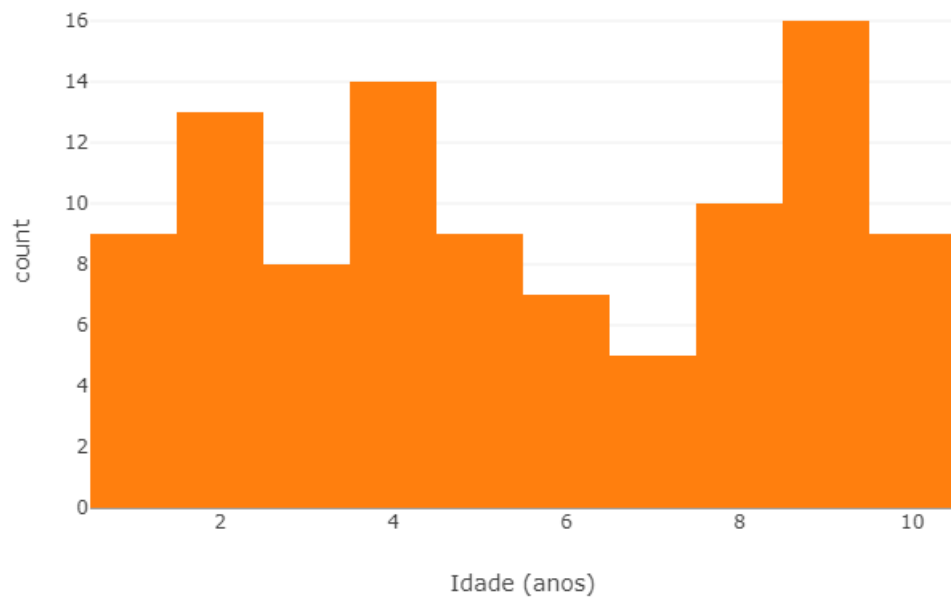
### Regressão Linear - Resultados



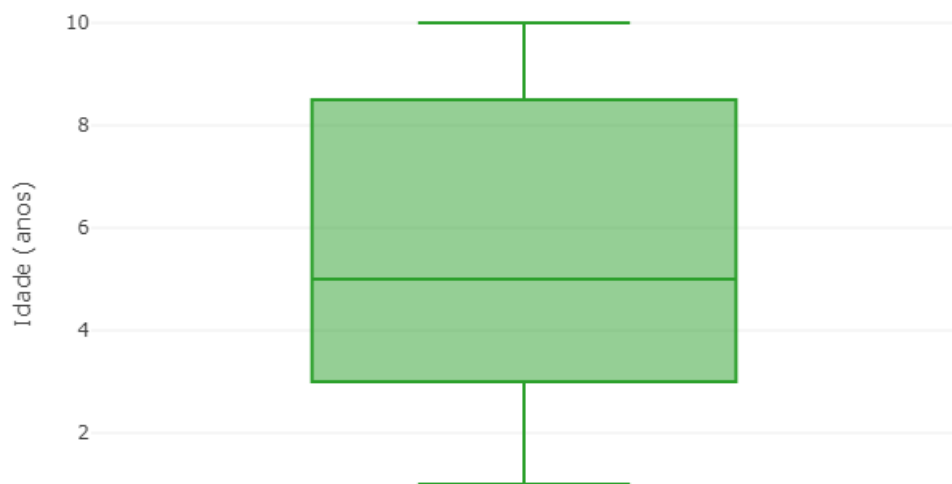
## Distribuição das Variáveis

A distribuição das variáveis foi explorada usando histogramas e boxplots. Estes gráficos ajudam a entender a dispersão e as características principais dos dados.

### Histograma da Coluna: Idade (anos)



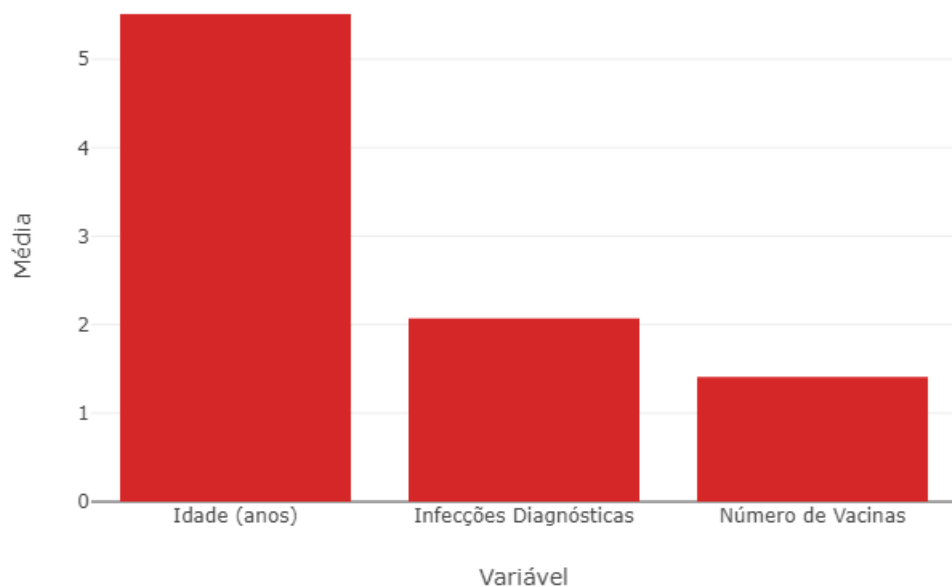
Boxplot das Variáveis Numéricas



## Média das Variáveis Numéricas

O gráfico de barras abaixo mostra a média das variáveis numéricas, o que ajuda a identificar tendências gerais nos dados.

Média das Variáveis Numéricas



### Projeto Integrador IV

desenvolvido por alunos do curso de Engenharia da Computação e Ciência de Dados

UNIVESP, 2º Semestre de 2024